DOCUMENT RESUME

ED 084 909 FL 004 533

AUTHOR Meeker, Mary; Meeker, Robert
TITLE Strategies for Assessing Intellectual Patterns in Black, Anglo, and Mexican-American Boys--or Any Other Children--and Implications for Education.
PUB DATE [73]
NOTE 32p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Anglo Americans; Aptitude Tests; Cognitive Tests; Cultural Factors; Culture Free Tests; *Educational Testing; *Group Intelligence Tests; Intelligence Quotient; *Intelligence Tests; Mexican Americans; *Minority Group Children; Negro Youth; Prognostic Tests; Spanish Speaking; Student Testing; *Test Bias; Test Construction; Test Interpretation; Test Reliability; Test Validity
IDENTIFIERS *Stanford Binet Intelligence Test

ABSTRACT

In this analysis of intelligence testing of minority group children, the implications of inadequate testing practices are discussed. Several aspects of test design are examined: deficiencies in intelligence testing, cultural bias, construct validity, and diagnostic utility. A sample set of results derived from a Stanford-Binet test administered to 257 respondents is examined: statistical data are included. The author concludes that "investigations of cultural biases in intelligence testing have established the fact that the most widely used test procedures are 'penalizing' for non-Anglo, lower socioeconomic groups." Teachers are cautioned of the dangers in using group-test results to plan academic programs geared to individual needs. (RL)

FILMED FROM BEST AVAILABLE COPY

ED 084909

Strategies for Assessing Intellectual Patterns in Black, Anglo, and Mexican-American Boys--or any other Children--and Implications for Education

Mary Meeker
Loyola-Marymount University, Los Angeles
and Robert Meeker
University of California, Los Angeles

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Today it is common practice for those who speak on behalf of disadvantaged students to be against IQ testing in the schools. The basis of their opposition is a familiar scenario: A student is tested; the test may be, for a variety of reasons, inaccurate; if so, the test results typically characterize the student as below normal; the characterization functions as a label which prompts attitudes and treatments that are subtly (or otherwise) transmitted to the student; the student, in turn, tends to reflect behavior that fulfills the expectation. In short, if the psychologist tells the teacher that Johnny is below average, the teacher will treat him as such, and Johnny will respond accordingly, all because the practice of testing is little related to the reason for testing.

No one doubts that this scenario has been, and still is, being played out countless times in schools throughout the nation. The situation, however, is not at issue; the

FL004 533

issue is what to do about it. To understand the issue better, we need to follow the scenario one step further.

Once the problem is acknowledged, a further dialogue ensues, in which there are four principals: those who represent the disadvantaged, those who are directly concerned with instruction, those who are involved in test construction, and, in the center of this dialogue, the school psychologist. Representatives of the disadvantaged are advocating the abolition of all IQ testing; psychometricians are analyzing the sources of test inaccuracies, and discussing means for making tests more valid, more reliable and culture-free; and, close at hand, the classroom teachers are requesting professional assistance for their daily encounters. The school psychologist must respond to all of them, and although it is the response to the classroom teacher that is critical, it is necessary to examine each response in turn.

To the representative of disadvantaged: The proposed abolition of IQ testing is well-intended, but ill-conceived; the abolition of all testing would only serve to drive the phenomenon underground. To deny teachers access to formal assessments is to force them to make informal assessments that are subject to the same sorts of deficiencies, which, not being exposed to scrutiny, are not open to easy identification and correction. (This also denies the need for formal assessments that some students must have to meet legislated qualifications in many special programs.)

To the psychometrician: Analyses of latent sources of testing deficiencies are helpful, and the proposed programs for rectification are welcomed, but operations cannot be suspended while we wait for more reliable and valid instruments. The job of constructing better tests is an exacting task that takes time, but the educational need is always immediate and on-going, so the best available instruments must be used.

The school psychologist is, then, in a position where he or she can neither forfeit nor defer responsibility. That does not mean doing business as usual; many of the cited injustices of the present practices can be eradicated by making some changes in the professional services provided. To best understand the nature and importance of these changes, we need to reexamine the dynamics of the opening scenario. The inescapable conclusion drawn from that scenario is that deficient assessments are almost uniformly insufficient, if not detrimental, for the ensuing instruction. It seems obvious that poor diagnosis would produce dysfunctional treatment, but we might look at the situation more closely, and rather than focus on the potentiality of error, we can ask why the effect of error is so overwhelming.

Upon reflection, one must wonder how a gross deficiency in assessment--and, by the nature of the case, the inaccuracy must be more than marginal--how an error of such magnitude can be perpetuated. Why, in other words, don't the child's strong abilities naturally assert themselves and thereby right

the situation? Consider the analogous situation in medicine: Doctors make diagnoses, and they are not always accurate, but hospitals are not full of healthy patients who have been mistakenly identified as sick.

The analogy is misleading, but instructive for its contrast; there are two essential points of difference: First, medical diagnosis is specific, and second, treatment is considered an extension of diagnosis in the sense that a patient's response to treatment is part of a continuous and reflexive process of diagnostic review. This latter, of course, does not happen in the schools. To underscore these points of contrast, we need only recast the opening scenario in medical terms: The doctor diagnoses a referral as merely sick or below average health, passes this global assessment on to the treatment personnel who, regarding the patient as "sick," put him under general hospital care (no prescribed treatment)--under these circumstances, the patient's general health may never show enough improvement to be discharged.

To bring this back to the main line of discussion, the problem is not the potentiality of error, so much as the fact that the assessment is so general in kind that, as a consequence, it bears little relation to treatment. If assessments are specific and prescriptively related to treatment, then there is considerably less chance that errors, when they do occur, will be perpetuated.

All of this serves as an extended prologue to considering the school psychologist's response to requests from classroom teachers for professional assistance. The response cannot be general, global, and unrelated to instructional treatment. In the study to follow we suggest one, though certainly not the only, means of being specific and prescriptive.

Deficiencies in Intelligence Testing

To provide a proper framework for our approach we need to look at the deficiencies of intelligence testing in the most general terms possible. If we ask the question, "What's wrong with intelligence testing?" we find researchers responding to three different aspects of the problem: (1) Cultural bias--tests (and test administration) are inadequate because they are predicated on a cultural norm that is penalizing to those outside the norm. (2) Construct validity--tests are inadequate because they systematically exclude important aspects of intelligence. (3) Diagnostic utility--tests are inadequate because they fail to provide adequate information for treatment. Those three aspects of the problem are clearly distinct. There could be, for instance, culture fair tests that were conceptually invalid, and there could be conceptually valid tests that were diagnostically sterile--any combination is possible, and, importantly, it is

generally conceded that the most widely used tests (and testing procedures) are, generally, deficient in all three respects.

Cultural Bias. The importance of cultural concomitants in interpreting intelligence test results is evident in a study by Mercer, et al. (1972). They found a direct relationship between cultural background and IQ test measures. Undifferentiated test results for Chicano and Black children were (on an average for both groups) about ten points below average. These undifferentiated results are similar to what other investigators have found; the importance of the Mercer project is that they could account for this below-average performance by "exogenous" cultural concomitants. Applying a five-factor index of "Anglicized" culture: (1) the mother wants the child to have an education beyond high school, (2) the parents are married, (3) the family are home owners, (4) the father has a skilled job, and (5) the family is relatively small and intact--the IQ scores were grouped according to the degree of "Anglized" cultural background. The differentiated average for the Anglo (score of 5) was average or slightly above average; on the other hand, members of the least Anglicized group (score of 0) were a standard deviation below the norm. This pattern of results held for both Blacks and Chicanos. The penalizing effects of a non-Anglo background are obvious and conclusive in this study.

The same general point has been underscored and amplified in many other studies. On a slightly different, but highly related line of investigation, a number of researchers (Pasamanick and Knoblock, 1955; Bloom, 1964; Bereiter, 1965; Gray and Klaus, 1965; Lesser, Fifer, Clark and others, 1965) have been concerned with the effects of examiner bias. Generally they have found that a cultural difference between examiner and examinee has an adverse effect on the resulting IQ score. The focus of these investigations is not blatant prejudice on the part of the examiner--that would be easily detected--but rather on the subtle effects of the rapport and language necessary for adequate responses for a power test of intelligence. In all, these studies underscore a significant problem in the procedures of traditional standardized intelligence testing.

Other investigators have been concerned with the nature of the intervening variables that might serve to explain why tests are culturally biased. In other words, granting the fact of cultural bias, a number of studies have concentrated on identifying the characteristic differences in test performances that might account for the cultural bias. Typical of this line of investigation is the longitudinal study of Hertzig, Birch, Thomas, and Mendez (1968). They "amplified" the normal mode of intelligence testing--in addition to the usual recording of right and wrong Binet answers, with each child using his preferred language, more detailed observations

of examinee responses were made. Each response was characterized as verbal or non-verbal, and further classified as to its elaboration (i.e., whether the response was limited to the expected one or was spontaneously extended or explained). They compared the responses from Puerto Rican children of lower-class blue-collar workers with responses from Anglo children of middle-class professionals. The Anglo children were significantly more verbal and elaborating in their responses. For those who know the response-dynamics of the testing situation, it is not unreasonable to conclude that the culturally related differences in mode of response would account, at least in some measure, for the culturally related differences in test performance. It is possible that children who are able to elaborate, even in a trial-and-error guess, have a better chance at arriving at an acceptable answer.

To summarize: Investigations of cultural biases in intelligence testing have established the fact that the most widely used tests and test procedures are "penalizing" for non-Anglo, lower socioeconomic groups.

Construct Validity. A second approach to the general issue of testing deficiencies is concerned with what intelligence tests are, in fact, measuring. In other words, this line of investigation questions whether the most widely used tests (especially the Binet, a power test, and the WISC, a speeded test of intelligence) are adequate as instruments of measurement. They obviously measure something, but is it the

whole, or even the most significant aspect of intelligence? While the construct validity of any testing instrument is always (in principle) open to question, the field of intelligence testing presented a situation where, for all practical purposes, IQ scores and intelligence had become (and are still considered by many to be) synonymous. A number of researchers have been concerned to break the "set" of a unidimensional, static concept of intelligence.

One line of investigation has been to question the assumed constancy of intelligence. (This research has, of course, been prompted, influenced, and guided by the work of developmental psychologists, preeminently Piaget.) As an example, the McCall, Hogarty, and Hurburt study (1972), at the Fels Research Institute, made a longitudinal study of general Binet IQ scores. Their investigations underscore the importance of the development aspects of intelligence, i.e., that a general index of intelligence does not hold constant for the same respondent over time. To quote their summary and conclusions:

> The most pronounced trend spanning the entire infancy period involved the manipulative exploration of objects that produced perceptual contingencies at 6 months, the imitation of simple fine motor and elementary verbal behavior particularly in a social contact at 12 months, verbal labeling and comprehension at 18 months, and verbal fluency and grammatical maturity at 24 months.
>
> Moreover, to label as "mental," performances at every age perpetuates the belief in a pervasive and developmentally constant intelligence. Consequently, the term mental as applied to infant behavior or tests should be abandoned in favor of some conceptually more neutral label, perhaps Piaget's "sensorimotor," "perceptual-motor," or even more specific classes of behaviors (e.g., exploration

of perceptual contingencies, imitation, language). The network of transitions between skills at one age and another is likely more specific and complex than once thought, and not accurately subsumed under one general concept.

Psychometricians have also questioned the (presumed) adequacy of a unidimensional index of intelligence. As early as the 1930's W. P. Alexander (1934), after Thurstone, found that general intelligence accounted for only 10% of success in shop achievement (spatial ability accounted for 13%, motivation for 48%, and 34% remained unaccounted). Research into specific intellectual abilities (as contrasted with general ability) has been developing ever since.

Most notable among these developments has been the work of Guilford and his associates (Guilford, 1956). Using factor-analytic techniques, they found sets of distinct intellectual abilities which could be conceptualized along three dimensions, which they referred to as the Structure of Intellect. (An elaboration of the theoretical SI model by Meeker was named the SOI for purposes of application; the schema is given later in this article.) Subsequent to this pioneering work, which used adult males as the subject population, other investigators (Meeker, 1963; Meyers, et al., 1964; Orpet and Meyers, 1966; Sitkei, 1966; Ball, 1972 [see her contribution elsewhere in this journal]) have found similar factors among normal, mentally retarded, physically handicapped, and gifted children.

The inadequacy of a general index of intelligence seems apparent and undoubtedly the trend toward greater differentiation will continue. Nonetheless, the instruments of general assessment will not be quickly or easily displaced in the school context for two reasons: First, the instruments are familiar to practitioners and they are, undeniably, statistically sound. Second, there is, at present, no practical substitute for the Binet and WISC; i.e., there are no differentiated abilities tests (group or individual) that can be used within the limits of time and personnel that are normally allocated to testing. In other words, general intelligence instruments, although inadequate, will find continued use as long as there are not practical specific-abilities tests available.

Diagnostic Utility. Diagnostic utility is, as the term implies, a practical consideration relating to a test's adequacy. Evaluations of utility are always made relative to some operational context. Obviously, in the present case, evaluation of diagnostic utility is being made with reference to the school context.

Two general points about diagnostic utility deserve comment. First, it is a legitimate concern. True, those who are theoreticians or pure researchers may not acknowledge the legitimacy of diagnostic utility as a criterion of test adequacy. They may make this judgment for themselves, but they

cannot presume to impose this judgment on those who use tests for diagnostic purposes. And, it would be obtuse for those who have diagnostic responsibility to disregard any test's diagnostic potential. Second, diagnostic utility should not be confused with predictive validity. A test's predictive validity is measured by its accuracy in predicting performance in non-test situations; a test's diagnostic utility is evaluated by its usefulness in prescribing effective treatment or intervention (such as, for example, reading tests which diagnose problem areas for the purpose of remediation). General intelligence tests have high predictive validity for school performance, but they are nearly useless as a basis for prescribing treatment. Generally, if a test is being used as a screening device, one looks for predictive validity; but if a test is being used as a guide for treatment, one looks for a test with diagnostic utility. The distinction is critical to the whole issue of intelligence testing as it relates to the disadvantaged; the fact that the tests, as currently used in the schools, have high predictive validity is, in a sense, the problem: as screening devices they work all too well; as diagnostic instruments they are, if left as is, actually dysfunctional.

For a test to have diagnostic utility it must be specifically and differentially related to treatments or interventions that are, practically, within the diagnostician's domain of control. Of the two criteria, the first needs little

elaboration. The more specific a diagnosis, the more specific the prescriptive treatment can be, and, consequently, the more exact the evaluation of the treatment process. The second point, the need for diagnostics to relate to the practical domain of control, deserves more elaboration. In the abstract it may seem obvious that if an instrument points to variables outside the diagnostician's domain of control, little practical use can be made of the information. Knowing that x-factor is related to y-ailment is useful for intervention only if x-factor can be controlled, manipulated, checked, or otherwise effected. For this reason, the diagnostic utility of SES-concomitant assessments would seem to be very limited; the fact that SES is a determinant of test performance leads nowhere in terms of direct prescriptive treatment since the socioeconomic status of the student is outside the domain of control for the school. (It may, of course, serve to caution the diagnostician not to take the test score at face value, but beyond that it provides little direction for treatment.)

The most widely used tests (the Binet and the WISC) have very limited diagnostic utility; as measures of general intelligence they offer little guidance for prescriptive treatment. As a practical and interim (until specific abilities tests can be developed*) remedy for this situation, Meeker (1963, 1969) has proposed a method for using Binet (or WISC) responses to derive differentiated assessments of

---

*Such a project is now in progress.

SOI abilities. This method has been used extensively in studies by Meeker (1965), Feldman (1970), Brown (1971), Karradenes (1971), Hays and Periera (1972), Hess (1972) and Manning (1972). The study report that follows is illustrative of the potential diagnostic utility afforded by differentiated indices of intelligence.

STUDY METHOD This study is based on item-response data from Stanford-Binet tests administered to 257 respondents. Using a technique described elsewhere (Meeker, M., 1963), the item responses were tallied according to the Structure-of-Intellect schema (see Fig. 1). All subjects were boys who resided in innercity Los Angeles urban communities.

- - - - - - - - - - - - - - -
(Insert Fig. 1 about here)
- - - - - - - - - - - - - - -

Respondents were from one of seven groups:

(1) MAS (4-5) Mexican-Americans, age 4 to 5, who took their tests in Spanish with a Mexican-American examiner.

(2) MAE (4-5) Mexican-Americans, age 4 to 5, who took their tests in English; they spoke English and their parents spoke English. An interpreter, when needed, was present in each examination.

(3) MAE (7-9) Mexican-Americans, age 7 to 9, who took their tests in English; they spoke English and their parents spoke English.*

---

*It was not possible to complete a sample of MAS (7-9) to contrast and compare with MAS (4-5).

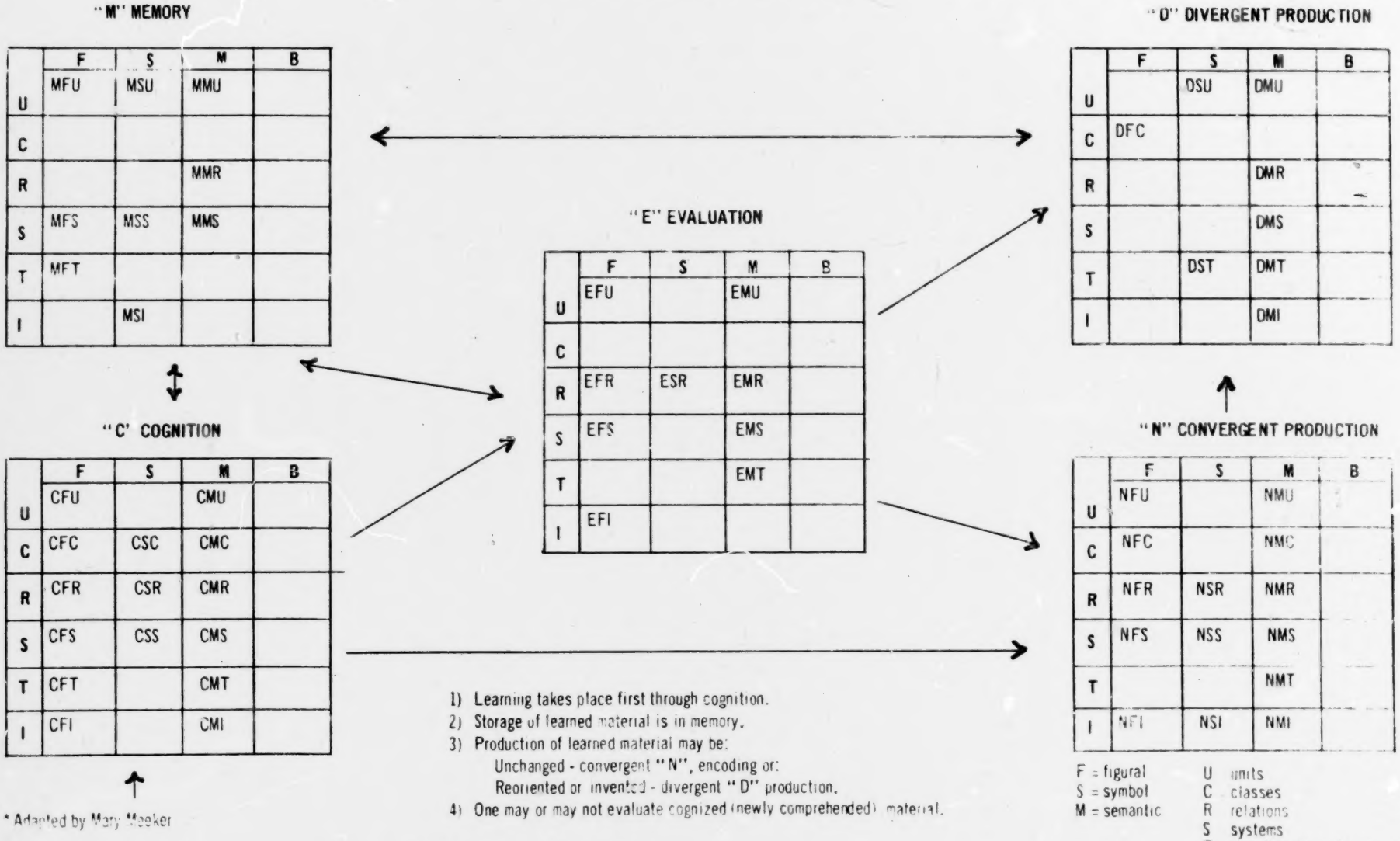
TALLY SHEET FOR STANFORD BINET (LM) TEMPLATES

Child's Name _____

I. Q. _____

STRUCTURE OF INTELLECT PROFILE

WITH A FLOW DIAGRAM OF THE PROCESSES*


"M" MEMORY
"D" DIVERGENT PRODUCTION
"E" EVALUATION
"C" COGNITION
"N" CONVERGENT PRODUCTION
1) Learning takes place first through cognition.
2) Storage of learned material is in memory.
3) Production of learned material may be:
Unchanged - convergent "N", encoding or:
Reoriented or invented - divergent "D" production.
4) One may or may not evaluate cognized (newly comprehended) material.
F = figural
S = symbol
M = semantic
U units
C classes
R relations
S systems
T transformation
I implication

	F	S	M	B
U	MFU	MSU	MMU	
C				
R			MMR	
S	MFS	MSS	MMS	
T	MFT			
I		MSI		

	F	S	M	B
U		DSU	DMU	
C	DFC			
R			DMR	
S			DMS	
T		DST	DMT	
I			DMI	

	F	S	M	B
U	EFU		EMU	
C				
R	EFR	ESR	EMR	
S	EFS		EMS	
T			EMT	
I	EFI			

	F	S	M	B
U	CFU		CMU	
C	CFC	CSC	CMC	
R	CFR	CSR	CMR	
S	CFS	CSS	CMS	
T	CFT		CMT	
I	CFI		CMI	

	F	S	M	B
U	NFU		NMU	
C	NFC		NMC	
R	NFR	NSR	NMR	
S	NFS	NSS	NMS	
T			NMT	
I	NFI	NSI	NMI	

* Adapted by Mary Meeker

Fig. 1. Structure of Intellect Profile.

(4) B (4-5) Blacks, age 4 to 5, tested in English by Black examiners.

(5) B (7-9) Blacks, age 7 to 9, tested by Black examiners.

(6) A (4-5) Anglos, age 4 to 5, tested by Anglo examiners.

(7) A (7-9) Anglos, age 7 to 9, tested by Anglo examiners.

Sample Description

	Age Range	IQ Range	IQ Mean	Sex	Number
MAS-4-5	4.9-5.9	79-113	90	M	37
MAE-4-5	4.9-5.9	77-123	95	M	33
MAE-7-9	7.0-9.11	76-144	101	M	35
BLACK-4-5	4.9-5.9	78-135	100	M	31
BLACK-7-9	7.0-9.11	77-153	103	M	24
ANGLO-4-5	4.11-5.9	80-132	101	M	33
ANGLO-7-9	7.0-9.11	79-145	104	M	64

One condition of the 4-5 year old sample was that none had had any formal preschool education; that is, none had been in Head Start, nursery, or coop preschool. It was our intent to try to get SOI-Binet profiles on the 4-5 year olds in an attempt to have a sample of entering kindergarteners who were "uncontaminated" by formal education.

We wanted to see what kinds of SOI abilities boys come to school with when they have had limited exposure to learning

of an academic nature. The reason for selecting the comparable age 7-9 group was to see what, if any, changes occurred in their SOI abilities due to exposure to traditional school learning.

The group identity of the respondents was retained in the tally of each item-response; as a result, each datum is characterized by a five-way classification:

GROUP--MAE(4-5), MAS(4-5), MAE(7-9), B(4-5), B(7-9), A(4-5), A(7-9).

OPERATION--Cognition, Memory, Evaluation, Convergent and Divergent Productions.

CONTENT--Figural, Symbolic, Semantic.

PRODUCT--Units, Classes, Relations, Systems, Transformations, Implications.

SCORE--Correct, Incorrect.

The five-way classification yields a potential data space of 1260 cells; the sampling distribution in the data space was too irregular to support a full multi-classification analysis, so each of the major SOI dimensions was analyzed independently (with consequent loss of information pertaining to between-dimension interactive effects). Multi-classification analyses for each of the SOI dimensions showed highly significant differences.

GROUP X OPERATION X SCORE $x^2 = 101.6457$ df = 24 $p<.0001$

GROUP X CONTENT X SCORE $x^2 = 154.1713$ df = 12 $p<.0001$

GROUP X PRODUCT X SCORE $x^2 = 170.044$ df = 30 $p<.0001$

Each of the above relationships was further analyzed with regard to the within-group and between-group effects. These results are of greatest interest for the present study since they afford two kinds of comparisons. The within-group analyses reveal general strengths and weaknesses profiles for each group, while the between-group analyses serve to anchor these evaluations in relationships to other groups, and, by implications, to the general population. In other words, if a group shows particular strength in, say, cognition (among the operations), that fact in itself would be helpful in planning instructional programs, and if, in addition, the group also shows strength in cognition in comparison with other groups, this would serve to reinforce the evaluation. Thus, in interpreting the results we look primarily to the within-group analyses since they are most useful for instructional prescriptions, and we look secondarily to between-group analyses as a means of anchoring the group ability profiles. Summaries of the within- and between-group analyses for each of the major dimensions of the SOI are presented in Tables 1, 2, and 3.

- -
Insert Tables 1, 2, and 3 about here
- -

Discussion. We offer this study as an illustration of the potential utility of specific ability assessment. Beyond that, we eschew group-oriented interpretations as generally

GROUPS	OPERATIONS				
	COGNITION between groups: P .003	MEMORY between broups: P .0000001	EVALUATION between groups: P .00000001	CONVERGENT between groups: P .279	DIVERGENT between groups: P .0006
Mexican-Americans (C.A.4-5) tested in spanish within group: P .000008		within: weak between: strong	within: strong between: strong		within: weak between weak
Mexican-Americans (C.A.4-5) tested in English within group: P .000000001	within: strong	within: weak between: weak	within: strong between: strong		
Mexican-Americans (C.A.7-9) tested in English within group: P .003	within: strong between: weak		within: weak between: weak	within: strong	within: weak between: weak
Blacks (C.A.4-5) within group: P .0000002	within: strong	within: weak between: weak			within: weak
Blacks (C.A.7-9) within groups: P .07	between: weak	between: strong	between: weak		
Anglo (C.A.4-5) within group: P .000000001	within: strong between: strong	within: weak between: weak			within: weak
Anglo (C.A.7-9) within group: P .000000001	within: strong	within: weak	within: weak between: weak		between: strong

	CONTENTS		
GROUPS	FIGURAL between groups: P .000000001	SYMBOLIC between groups: P .00000002	SEMANTIC between groups: P .002
MA-5 4-5 within groups: P<.000000001	within: strong between: strong	within: weak between: weak	within: weak between: weak
MA-E 4-5 within groups: P<.00000001	within: strong between: strong	within: weak between: weak	within: weak
MA-E 7-9 within groups: P<.01	within: strong between: weak	within: weak	between: weak
Blacks 4-5 within groups: P<.000000001	within: strong between: weak	within: weak between: weak	
Blacks 7-9 within groups: P<.06	between: weak	between: strong	
Anglos 4-5 within groups: P<.000000001	within: strong	within: weak between: weak	
Anglos 7-9 within groups: P<.000000001	within: strong between: weak	within: weak between: strong	between: strong

	PRODUCTS					
GROUPS	UNITS between groups: P .0000002	CLASSES between groups: P .09	RELATIONS between groups: P .06	SYSTEMS between groups: P .000000001	TRANSFORMATIONS between groups: P .00002	IMPLICATIONS between groups: P .0005
MA-S 4-5 within groups: P .000008	within: strong between: strong	within: strong	within: weak	within: strong between: strong	within: weak	within: weak between: weak
MA-E 4-5 within groups: P .000000001	within: strong between: strong	within: strong		between: strong	within: weak between: weak	within: weak between: weak
MA-E 7-9 within groups: P .003	between: weak	within: strong		within: weak between: weak		within: strong
Blacks 4-5 within groups: P .0000002	within: strong	within: strong			within: weak between: weak	
Blacks 7-9 within groups: P .07					between: strong	
Anglos 4-5 within groups: P .000000001		within: strong		within: weak	within: weak	
Anglos 7-9 within groups: P .000000001	between: weak	within: strong		within: weak between: weak	within: strong between: strong	within: strong between: strong

Within Groups: Comparisons of Contributions to Chi Square

Disadvantaged Group	OPERATIONS					CONTENTS			PRODUCTS					
	Cognition	Memory	Evaluation	Convergent	Divergent	Figural	Symbolic	Semantic	Units	Classes	Relations	Systems	Transformations	Implications
MA-S 4-5		weak	strong		weak	strong	weak	weak	S	S	W	S	W	W
ME 4-5	strong	weak	strong			strong	weak	weak	S	S			W	W
ME 7-9	strong		weak	strong	weak	strong	weak			S		W		S
Blacks 4-5	strong	weak			weak	strong	weak		S	S			W	
Blacks 7-9														
Anglos 4-5	strong	weak			weak	strong	weak			S		W	W	
Anglos 7-9	strong	weak	weak			strong	weak			S		W	S	S

Between Groups: Comparisons of Contributions to Chi Square

	OPERATIONS					CONTENTS			PRODUCTS					
	Cognition	Memory	Evaluation	Convergent Production	Divergent Production	Figural	Symbolic	Semantic	Units	Classes	Relations	Systems	Transformations	Implications
MA-S 4-5		S	S		W	S	W	W	S			S		W
MA-E 4-5		W	S			S	W		S			S	W	W
MA-E 7-9	W		W		W	W		W	W			W		
Blacks 4-5		W				W	W						W	
Blacks 7-9	W	S	W			W	S						S	
Anglos 4-5	S	W					W							
Anglos 7-9			W		S	W	S	S	W			W	S	S

S = Strong
W = Weak

LEGEND	Tested in	Age
MA-S 4-5 = Mexican Americans	Spanish	4-5
MA-E 4-5 = Mexican Americans	English	4-5
MA-E 7-9 = Mexican Americans	English	7-9
Blacks 4-5 = Blacks		4-5
Blacks 7-9 = Blacks		7-9
Anglos 4-5 = Anglos		4-5
Anglos 7-9 = Anglos		7-9

dysfunctional for educational practice. While group results might have limited utility for general instructional planning, it should be patently obvious that an individual student's profile of abilities on any or all of the SOI dimensions may be vastly different from his group's profile on any or all of the SOI dimensions. As obvious as this may be statistically, one nonetheless finds, in instructional practice, that group-type diagnoses are used as bases for prescribing individual treatment. We explicitly disown any such use that might be made of the results; indeed, the larger point at issue--that specific, treatment-related, individual assessment is an immediate remedy for intelligence testing abuses--would be subverted by using group-oriented data as a substitute for individual diagnostics.

GROUPS	OPERATIONS					Within Group (between operation probabilities
	COGNITION	MEMORY	EVALUATION	CONVERGENT PRODUCTION	DIVERGENT PRODUCTION	
Mexican-Americans (CA 4-5) Tested in Spanish	522 / 392	96 / 103 L^W H^b	235 / 163 H^W H^b	137 / 107	38 / 75 L^W L^b	$p < .001$
Mexican-Americans (CA 4-5) Tested in English	336 / 242 H^W	55 / 127 L^W L^b	146 / 105 H^W H^b	124 / 137	45 / 48	$p < .001$
Mexican-Americans (CA 7-9) Tested in English	148 / 164 H^W L^b	91 /134	59 / 109 L^W L^b	89 / 85 H^W	34 / 69 L^W L^b	$p < .003$
Blacks (CA 4-5)	380 / 293 H^W	81 / 148 L^W L^b	140 / 169	134 / 157	46 / 66 L^W	$p < .001$
Blacks (CA 7-9)	278 / 270 L^b	201 / 173 H^b	140 / 186 L^b	172 / 182	95 / 94	$p < .07$
Anglos (CA 4-5)	391 / 274 H^W H^b	86 / 181 L^W L^b	155 / 166	166 / 164	55 / 75 L^W	$p < .001$
Anglos (CA 7-9)	754 / 566 H^W	353 / 514 L^W	362 / 435 L^W L^b	389 / 361	224 / 206 H^b	$p < .001$
Between Groups (within operations) probabilities	$p < .003$	$p < .001$	$p < .001$	$p < .28$	$p < .001$	

H^W--high within group
H^b--high between groups
L^W--low within group
L^b--low between groups

Table I. Aggregate Group Scores for Operations Categories (items passed / items failed)

GROUPS	CONTENTS: FIGURAL	SYMBOLIC	SEMANTIC	Within Group (between contents) probabilities
Mexican-Americans (CA 4-5) Tested in Spanish	566 / 220 H^W H^b	30 / 91 L^W L^b	432 / 529 L^W L^b	$p < .001$
Mexican-Americans (CA 4-5) Tested in English	313 / 153 H^W H^b	23 / 84 L^W L^b	370 / 422 L^W	$p < .001$
Mexican-Americans (CA 7-9) Tested in English	102 / 94 H^W L^b	66 / 111 L^W	253 / 356 L^b	$p < .01$
Blacks CA 4-5)	313 / 236 H^W L^b	51 / 117 L^W L^b	417 / 480	$p < .001$
Blacks (CA 7-9)	190 / 154 L^b	145 / 150 H^b	551 / 601	$p < .06$
Anglos (CA 4-5)	318 / 211 H^W	49 / 123 L^W L^b	489 / 526	$p < .001$
Anglos (CA 7-9)	495 / 386 H^W L^b	255 / 387 L^W H^b	1332 / 1309 H^b	$p < .001$
Between Groups (within Contents) probabilities	$p < .001$	$p < .001$	$p < .002$	

H^W--high within group
H^b--high between groups
L^W--low within group
L^b--low between groups

Table 2. Aggregate Group Scores for Content Categories (items passed/ items failed)

PRODUCTS

GROUPS	UNITS	CLASSES	RELATIONS	SYSTEMS	TRANSFORMATIONS	IMPLICATIONS	Within Group (between products probabilities
Mexican-Americans (CA 4-5) Tested in Spanish	295/152 H^W H^b	137/59 H^W	168/187 L^W	179/114 H^W H^b	108/123 L^W	141/205 L^W L^b	$p < .001$
Mexican-Americans (CA 4-5) Tested in English	168/120 H^W H^b	82/38 H^W	140/132	111/103 H^b	72/98 L^W L^b	133/168 L^W L^b	$p < .001$
Mexican-Americans (CA 7-9) Tested in English	97/125 L^b	29/26 H^W	79/123	45/108 L^W L^b	56/63	115/116 H^W	$p < .003$
Blacks (CA 4-5)	181/158 H^W	83/47 H^W	159/149	117/145	77/138 L^W L^b	164/196	$p < .001$
Blacks (CA 7-9)	215/191	48/37	197/209	146/188	98/87 H^b	182/193	$p < .07$
Anglos (CA 4-5)	186/176	90/44 H^W	182/159	109/148 L^W	94/120 L^W	195/213	$p < .001$
Anglos (CA 7-9)	484/463 L^b	133/89 H^W	453/475	250/434 L^W L^b	251/189 H^W H^b	511/432 H^W H^b	$p < .001$
Between Groups (within products) probabilities	$p < .001$	$p < .09$	$p < .06$	$p < .001$	$p < .001$		$p < .001$

H^W--high within group
H^b--high between groups
L^W--low within group
L^b--low between groups

Table 3. Aggregate Group Scores for Products Categories (items passed / items failed)

References

Alexander, W. P. Research in guidance. A theoretical basis. Occupations, April 1934, 12, 75-91.

Baer, J. How delinquents think. Unpublished master's thesis, Wayne State University, 1970, pp. 5, 7.

Ball, R. S. Comparison of thinking abilities of 5-year-old white and black children in relation to certain environmental factors. Proj. 9-70-0067. Tempe, Ariz.: Arizona State University, May 1972.

Ball, R. S., & Merrifield, P. R. A longitudinal assessment of thinking of pre-literate children during a two-year period. Tempe, Ariz.: Arizona State University, 1971.

Bardon, J. I. The baby and the bath water: The case for discrimination between use and abuse. The Journal of Special Education, 5, No. 1.

Bereiter, C., et al. An academically oriented preschool for culturally deprived children. Paper presented at meeting of the American Educational Research Association, Chicago, Illinois, February 1965.

Bloom, B. S. Stability and change in human characteristics. New York: Wiley, 1964.

Bortner, M. Phrenology, localization, and learning disabilities. The Journal of Special Education, 5, No. 1.

Brown, D. L. Variations in test response of preschool children by sex and socioeconomic level related to Guilford's structure of intellect. Unpublished doctoral dissertation, University of Pittsburgh, 1971.

Cohen, J. The factorial structure of the WISC at ages 7-6, 10-6, and 13-6. Journal of Consulting Psychology, 1959, 23, 285-299.

Contreras, L. Comparison of group performance on SOI factors of retarded Mexican-Americans. Unpublished master's thesis, Long Beach State University, 1973.

Deutch, M. Facilitating development in the preschool child: Social and psychological perspectives. Merrill-Palmer Quarterly, 1964, 10, 249-263.

Deutch, M., & Brown, B. Social influences in Negro-white intelligence differences. Journal of Social Issues, 1964, 20, 24-35.

Dreger, R. M., & Miller, K. S. Comparative psychological studies of Negroes and whites in the United States. Psychological Bulletin, 1960, 57, 361-402.

Feldman, B. Prediction of first-grade reading achievement from selected structure-of-intellect factors. Unpublished doctoral dissertation, University of Southern California, 1970.

Garcia, J. IQ--The conspiracy. Psychology Today, September 1972.

Gray, S. W., & Klaus, R. A. An experimental preschool program for culturally deprived children. Child Development, 1965, 36, 887-898.

Guilford, J. P. The structure of intellect. Psychological Bulletin, 1956, 52, 267-293.

Guilford, J. P. The nature of human intelligence. New York: McGraw-Hill, 1967.

Havighurst, R. J., & Moorefield, T. E. The nature and needs of the disadvantaged. The Sixty-sixth Yearbook of the National Society for the Study of Education. Chicago: The University of Chicago Press, 1967.

Hays, B. H., & Pereira, E. R. Effect of visual memory training on reading ability of kindergarten and first grade children. The Journal of Experimental Education, 1972, 41, No. 1.

Hertzig, M. E., Birch, H. G., Thomas, A., & Mendez, O. A. Class and ethnic differences in the responsiveness of preschool children to cognitive demands. Monograph of the Society for Research in Child Development, 1968, 33, 1, Ser. 117.

Jacobs, P. I., & Vandeventer, M. Evaluating the teaching of intelligence. Educational and Psychological Measurement, 1972, 32, 235-248.

Karadenes, M. A comparison of differences in achievement and learning abilities between Anglo and Mexican-American

children when the two groups are equated by intelligence. Unpublished doctoral dissertation, The University of Virginia, 1971.

Lesser, G. S., Fifer, G., & Clark, D. H. Mental abilities of children from different social class and cultural groups. Monograph of the Society for Research in Child Development, 1965, 30, 4, Serial No. 102.

Lesser, G. S., & Stodolsky, S. S. Learning patterns in the disadvantaged. Harvard Educational Review, Reprint No. 6, 1967, 546-593.

Lovinger, R. J., Harris, A. J., & Cohen, J. Factor analysis of intellectual performance in disadvantaged Negro adolescents. Paper presented at the meeting of the American Educational Research Association, Chicago, February 1966.

McCall, Hogarty, and Hurburt. The McCall, Hogarty, and Hurburt Study, 1972.

Mallory, W. A. Abilities and developmental changes in elaborative strategies in paired-associate learning of young children. Journal of Educational Psychology, 1972, 63, 3, 202-217.

Manning, E. Teaching divergent thinking to gifted children. Title III Project. Whittier, Calif.: East Whittier Schools, 1972.

Meeker, M. N. The NSMP behavior samplings of the Binet. Paper presented at the meeting of the American Psychological Association, 1963, published by the Los Angeles County Schools, 1964.

Meeker, M. N. A procedure for relating Stanford-Binet behavior samplings to Guilford's structure of the intellect. Journal of School Psychology, 1965, 3, 26-36.

Meeker, M. N. The structure of intellect: Its interpretation and uses. Columbus, Ohio: Charles E. Merrill Publishing Co., 1969.

Meeker, M. N., & Sexton, K. M., CSJ. Structure-of-intellect abilities workbook, 2nd edition. Los Angeles: Loyola University of Los Angeles, 1971.

Mercer, J. IQ--the lethal label. Psychology Today, September 1972.

Meyers, C. E., Dingman, H. F., Orpet, R. E. Four ability-factor hypotheses at three preliterate levels in normal and retarded children. Monograph of the Society for Research in Child Development, 1964, 29, 5.

Nelson, L. L., & Kogan, S. Competition the star spangled scr[illegible]ble. Psychology Today, September 1972.

Newland, T. E. Sp[illegible] education--instrument of change in education [illegible] the '70s. Selected Papers from the University [illegible] Virginia Lecture Series, 1970-71.

Orpet, R. E., & Meyers, C. E. Six structure-of-intellect hypotheses in six-year-old children. Journal of Educational Psychology, 1966, 57, 341-346.

Pasamanick, B., & Knobloch, H. The contribution of some organic factors to school retardation in Negro children. Journal of Negro Education, 1958, 27, 4-9.

Pasamanick, B., & Knobloch, H. Early language behavior in Negro children and the testing of intelligence. Journal of Abnormal and Social Psychology, 1955, 50, 401-402.

Scagliotta, E. G. And the pendulum swings. The Journal of Special Education, 5, No. 1.

Sitkei, G. Comparative structure of intellect in middle and lower class four-year-old children in two ethnic groups. Unpublished doctoral dissertation, University of Southern California, 1966.

Stier, J. H. Report Study No. 1...Average daily attendance summary. Mimeographed paper, Santa Monica Unified School District, Santa Monica, California, 1969-1970, p. 1.

Stott, L. H., & Ball, R. S. Infant and preschool mental tests: Review and evaluation. Monograph of the Society for Research in Child Development, 1965, 30, 3.

Sutcliffe, J. P. A general method of analysis of frequency data for multiple classification designs. Psychological Bulletin, 1957, 54, No. 2.

Watson, P. IQ the racial gap. Psychology Today, September 1972.